

No Child Left Behind: The Mathematics of Guaranteed Failure

Lowell C. Rose

I. History and Background of NCLB

The signing of the No Child Left Behind Act on January 8, 2002, moved the federal effort to influence K–12 schooling to a new and higher level—more aggressive, focused, and directive. The act requires that school districts and schools demonstrate adequate yearly progress (AYP) toward a particular goal: universal student achievement of standards established by each state. Each year, school districts and schools that do not achieve AYP will be labeled “Did Not Make AYP,” and after two such years they may suffer sanctions that include loss of federal funding, termination of staff, and dissolving the school district.

The new law was accompanied by promises of higher funding, both to enhance the prospects of success and to serve as an incentive for compliance. However, two years after NCLB was signed amid bipartisan euphoria and more than one year into implementation, its chances of success and prospects for increased funding are, at best, uncertain. In Indiana, for example, Title I funding has increased only 11 percent over the year before NCLB was signed; virtually all of that amount funds new programs mandated by NCLB; and members of the state’s congressional delegation say that there is little prospect of increased funding in future years. Many educators suspect that NCLB may join special education as a major unfunded mandate.

The fanfare that accompanied NCLB seemed to herald a new federal initiative. Technically, however, NCLB is but the latest reauthorization of the 1965 Elementary and Secondary Education Act. During the past thirty-seven years, it became the practice to assign each reauthorization a distinctive title: hence, 2001’s “No Child Left Behind Act.” Each version of the ESEA has sought in particular to improve the achievement of low-performing students. Over the years an increasing emphasis on out-

comes culminated in the 1994 reauthorization (dubbed “Goals 2000”), which emphasized higher standards, testing based on those standards, and demands for accountability. Also along the way, federal frustration over a perceived lack of progress began to grow.

That was the background against which the provisions of the 2001 NCLB reauthorization were framed. Both presidential candidates in 2000 had promised change and improvement in education. Upon taking office, the new Bush administration garnered support by focusing on the seductive promise that no child would be left behind. Those heady days of bipartisanship in education were symbolized by pictures of the president and a partisan archrival, Sen. Edward Kennedy, dining at the White House to discuss the future of K–12 education.

II. The Problem with NCLB

So just what is wrong with NCLB? To understand the answer, it’s important to understand one of the act’s major requirements (and greatest strengths): breaking out test scores by demographic groups—black, Hispanic, Native American, Asian, special education, LEP (limited English proficient), economically disadvantaged, and white. The act requires demographic breakouts in order to avoid a classic anomaly in education data: satisfactory overall group performance that simultaneously obscures poor performance by subgroups. Unreported subgroup performance is one reason U.S. society and policymakers have tolerated achievement gaps between certain subgroups—students living in poverty, blacks, Hispanics—and other groups, usually white or more advantaged students.

The strength of NCLB is in forcing educators to confront disparate student achievement. That strength is, however, negated by the act’s single numerical goal for all groups and its requirement that each group reach 100 percent proficiency in twelve years, no matter the achievement level at which it began. NCLB’s single goal for all breakout groups, applied without concern for where a group starts or how much improvement it demonstrates, virtually guarantees immediate failure for school districts and schools that enroll high percentages of the most challenged students. (See Part IV, “Calculating Your District’s AYP.”)

The single-goal approach also guarantees that students who have had the greatest difficulty achieving must demonstrate the greatest progress. Virtually all the research on aspiration and student achievement has found, however, that improvement must be measured against the point at which the student begins; that it is hard work; that it comes unevenly, with significant gains accompanied by plateaus and temporary setbacks; and that improvement requires ongoing effort and commit-

ment. The literature as well as direct professional experience and common sense tell us that both effort levels and chances of success are enhanced by goals that force stretching but are nonetheless, given enough effort and commitment, within reach.

Unreported subgroup performance is one reason U.S. society and policymakers have tolerated achievement gaps between certain subgroups—students living in poverty, blacks, Hispanics—and other groups.

The NCLB scheme, ignoring the research consensus, assumes that improvement is continuous and consistent and that goals can be reached in a fixed time, regardless of the distance to be traveled. It says, simply, that we use 2001–2002 as the base year and progress in equal annual steps for twelve years, when 100 percent proficiency will be achieved. At the risk of oversimplifying, *measuring improvement from the point at which the school district, school, or breakout performance begins, using measures based on same-student performance*, would avert the NCLB's harshest consequences almost immediately. The school-improvement plans of Indiana and several other states already employ such an improvement focus.

However, even with realistic performance goals and measurements based on improvement by same-student cohorts, the issue of what constitutes realistic achievement will have to be addressed. This writer has yet to encounter a single person who believes that 100 percent proficiency for all students is a realistic or reachable goal.

III. Applying NCLB to Indiana: A Hypothetical

A good place to examine the potential impact of NCLB is in my home state of Indiana, one of the earliest states to embrace the accountability movement. Structured efforts to improve student achievement in Indiana began in the mid- to late 1980s. Indiana became one of seventeen states that complied fully with the provisions of the 1994 initiative, Goals 2000; when NCLB was signed in 2002, Indiana had already been measuring its own AYP benchmarks for four years. In addition, Indiana's Public Law 221, passed in 1999, places schools in categories based on improvement shown, with the first such placements to be made after the fall 2005 testing. Under P.L. 221, student progress for grades 3 to 10 will be assessed through same-student comparisons that use year-to-year test scores. In 2000, a graduation-qualifying exam at grade 10 was implemented, with few of the problems encountered in other states.

The passage of NCLB provided good news immediately for Indiana's budgetary woes: the funding needed to support the planned expansion

of ISTEP (Indiana Statewide Testing for Educational Progress) testing. The Indiana Department of Education (IDOE) accepted the challenge of NCLB and began planning to implement NCLB. The IDOE planning effort involved all those interested in K–12 schools (including the author, as a consultant to the Indiana Urban Schools Association [IUSA]). As a result, Indiana became one of five states placed on the fast track for federal approval.

As familiarity with the demands of NCLB grew, however, the law's ultimate feasibility came into question in many quarters. During that period, I developed and made several presentations on implementing NCLB in Indiana. In the process I used the results of the ISTEP tests given in September 2001 to calculate AYP for the participating schools (at that point, Indiana's AYP determinations for the NCLB base year, 2001–2002, were not yet made). I planned to use the figures to provide specific examples of the way AYP data could be used to focus instruction. Soon, however, I realized that NCLB's AYP calculations doomed the vast majority of schools and virtually every school district in Indiana to failure, with the schools that serve the highest percentage of challenged students achieving failure first.

In general, I concluded that applying NCLB to Indiana using 2001 data without mitigating measures would produce drastic and unexpected results. Here are the projections I made for Indiana school districts and schools based on the 2001 test data:

- Two hundred sixty-nine of the state's 293 school districts would have failed to achieve AYP.
- The twenty-four school districts that would have achieved AYP had an average enrollment of 770 students and an average minority enrollment of 1.7 percent.
- Each of the thirty school districts in the Indiana Urban Schools Association would have failed to make AYP. (The thirty-four IUSA school districts tested 93.8 percent of the state's black students, 68.2 percent of the Hispanic students, and 56.4 percent of the free- or reduced-lunch students.)
- Sixty-eight percent of the schools in the IUSA would have failed to achieve AYP.
- The failing schools in the IUSA would have included 95 percent of the high schools, 92 percent of the middle schools, and 57 percent of the elementary schools.

Those results reflect a particularly pernicious consequence of the way AYP is to be calculated. As mentioned earlier, larger schools test more students, which means more breakouts and a greater chance of fail-

ure to achieve AYP. Larger numbers tested also mean that less relief will be gained from applying the test of statistical significance, a test used to guarantee that differences are real. In addition, NCLB applies a single goal without concern for where a group starts or how much improvement it demonstrates. Therefore, when diversity adds more students who start far from the goal, the odds of achieving AYP diminish. At the school district level, achieving AYP in Indiana is almost beyond reach.

To see why it is vital to take into account both the starting point and the ability of the students in a breakout group, let's examine how the act would treat Indiana's special education students. The data for special education students in Indiana measured in the 2001 ISTEP testing showed that 252 of the 256 breakout groups would have failed to achieve AYP. (Communities with high socioeconomic levels furnished the four breakout groups that would have achieved AYP, which leads me to speculate that the basis of their special education placement was something other than cognitive ability.)

Most of Indiana's special education breakouts in 2001 would have missed NCLB targets by twenty to forty percentage points. Even taking into account NCLB's alternative assessments for some special education students, it appears that, regardless of the effort put forth, most special education breakout groups will not meet NCLB's standards. Special education groups will bear the brunt of the failure of school districts and schools to make AYP. It is difficult to believe that such a result is good for either special education or the school-improvement effort.

Given such projections, the logical question becomes: "Will any purpose be served by giving the schools that enroll a high percentage of disadvantaged and minority students a label that will be equated with 'failing'?" After all, Indiana's recent scores on the National Assessment of Educational Progress (NAEP) and other external measures suggest that the state's students have performed well in comparison to students in other states and consistently improved in recent years.

I realized that NCLB's AYP calculations doomed the vast majority of schools and virtually every school district in Indiana to failure.

But beyond that, if the "retro" exercise in applying NCLB to Indiana's 2001 numbers proves anywhere close to accurate, NCLB's methodology guarantees that matters will progressively worsen. For reasons related to the state of technology, more grades will be added to the testing pool each year, so "Did Not Make AYP"—which will be understood by everyone to mean "This school district or school is a failure"—will occur more often each year. The handful of school districts and schools left standing in 2005 will then be further challenged when, in compliance with NCLB,

the goals will jump by 25 percent of the difference between the initial goals and 100 percent proficiency.

Due in part to abbreviated media coverage, few people realize that AYP determinations and the possibility of improvement and corrective action status apply not only to schools but to school districts. AYP and NCLB sanctions may, in fact, prove most severe and immediate at the school district level. A district that fails to achieve AYP for two consecutive years will move to “school improvement” status. It must, at that point, work with the state to develop a plan for achieving AYP. A school district failing to make AYP for the fourth consecutive year moves to “corrective action” status, which under NCLB authorizes the state to exercise one of seven options:

- Reduce programmatic or administrative funds.
- Replace the curriculum.
- Terminate personnel relevant to the failure.
- Move some schools from the jurisdiction and provide alternative governance.
- Appoint a receiver to replace the superintendent and the board.
- Abolish and restructure the school district.
- Provide choice for school district students to attend successful neighboring school districts.

How does Indiana plan to proceed in relation to the seven options? The indication to date has been that none is permissible under Indiana law and that virtually all the state will be able to do is consult with school districts on how they spend their money. Given the federal government’s supremacy in distributing money, that statement is of little comfort.

In early 2003 I spoke to twenty relatively small school districts. Six of those districts will undoubtedly fail to make AYP—even though every school in them will achieve AYP. Think of that in terms of the purposes for which the school district exists!

IV. Calculating Your District’s AYP

Judge NCLB and its AYP system not in the abstract or based on what others say. To really understand AYP, familiarize yourself with the process. Take actual data for school districts and schools with which you are familiar, and calculate AYP.

I have calculated AYP several thousand times for school districts and schools with large enrollments, small enrollments, great diversity, and little or no diversity. That hands-on experience has led me to conclude that the NCLB’s AYP system will prove useless for improving student achieve-

ment. AYP targets not individual students but school districts and schools, and it effectively guarantees that virtually all of them will be labeled failures. That inevitability derives from the mandated mechanics of AYP.

The mechanics of AYP. The overall group and each breakout group in a school district or school are required to meet goals specifying the percentage of students who must pass the state tests on English and math. The goals are fixed by the state using a formula prescribed in NCLB. The goal is the same for all groups.

To determine if your school achieved AYP, you start with the overall group and ask, “Did the group make the goal in English?” If the answer is yes, you ask the same question for math. If the answer again is yes, you have two yeses. Now move to the first breakout group. A breakout group is any group—say, students with disabilities—for which the number of students tested in the school district or school is greater than the threshold number (the “N”) stated in the state plan. In Indiana and most other states the N is 30. Ask the same two questions for the first breakout and then repeat the process for each of the other breakouts. The first “no” answer means that the school district or school did not make AYP.*

A predominantly white elementary school with an enrollment of 250 or less will seldom miss achieving AYP because it won't contain groups large enough.

The typical urban school district in Indiana includes between four and seven breakout groups. The more breakouts, the more difficult it becomes to make AYP. In a school district or school with every one of the eight breakout groups, it will be necessary to answer “yes” eighteen times (twice for the eight breakout groups, and twice more for the overall group) to meet AYP based on test data alone. But for those with all yeses on the questions related to test data, there are still hurdles to jump. To protect against the possibility that school districts or schools might hold low-achieving students out of the testing, NCLB’s authors included a requirement that 95 percent of the overall group and 95 percent of all students in each breakout group must be tested in English and 95 percent in math. The N required for a breakout differs from state to state. In Indiana, it is thirty students tested. Apply that standard next. That adds

*In calculating AYP, consider the compound probabilities alone: if every group has a 99 percent chance of making AYP—nearly ideal conditions that are rarely seen—the odds that all thirty-seven groups will achieve AYP are equal to .99 to the 37th power, which is 0.689, or only slightly more than two out of three. Thus, despite the fact that our hypothetical school is nearly ideal, there is a nearly one in three chance that it will not achieve AYP. If each breakout has an 80 percent chance of making AYP—probably not bad in the real world—the chances that all thirty-seven will achieve AYP are fewer than 3 in 10,000. If one does not consider that situation impossible, factor in that the reason for defining a breakout group in the first place is usually to focus on students who have historically underachieved.—Ed.

two more required yeses for the overall group and two for each breakout group, bringing the number to thirty-six for a school district or school with the full complement of breakouts.

NCLB also requires at least one secondary indicator for each school. At the high school level, it must be the graduation rate. For other schools, the secondary indicator is determined by the state. Indiana's goal is a graduation rate of 95 percent for its high schools and 95 percent average daily attendance for all other schools. The secondary indicator, at least in Indiana, applies only to the overall group, thereby bringing the possible number of yeses required to thirty-seven.

It is plain to see how small schools with little diversity can achieve AYP. White breakouts achieve AYP. A predominantly white elementary school with an enrollment of 250 or less will seldom miss achieving AYP because it won't contain groups large enough. Given the N of 30 used in most states, it will have few, if any, breakouts. Case in point: the principal of an elementary school studied the data, realized his 600-plus-student elementary school would not make AYP, and called me to say, "If we split this school into three schools in the same building, we wouldn't have any breakouts and all three would make AYP."

Extend the game as long as you can. It is an interesting, albeit perverse, exercise to see how far any school or district can "drill down" into its breakout groups before it fails. The first rule in extending the game is to avoid beginning with the special education breakout. Most school districts and large schools have such breakouts and few will make AYP. As already noted, using the 2001 data, 256 of Indiana's 293 school districts would have had special education breakouts in the 2001 testing, and all but three would have failed to achieve AYP. I used the more-recent 2002 ISTEP data to calculate AYP for the thirty-four school districts in the Indiana Urban Schools Association. Had I started with the special education breakouts, I would have identified each of the school districts as "Did Not Make AYP." The game would have been over.

Absent significant change, NCLB will at minimum fail to improve schooling and do nothing to aid low-performing students.

To extend the game one must carefully choose the order in which the breakout groups are taken. If Indiana is a typical state, the breakout for Asians is a good place to start. That group is likely to score highest in terms of the goals. The white breakout also will earn two yeses in most schools. From that point on, the prospects of making AYP are iffy; nonetheless, there is much to be learned from extending the game. For example, free/reduced-lunch, Hispanic, black, Native American, and LEP

groups are among the least likely to make AYP. The finding should not be surprising given what we know about the achievement gap.

Drilling down further, however, you will quickly conclude from the free/reduced-lunch percentages that there is legal poverty and then there is real poverty. Free/reduced-lunch groups in urban areas are unlikely to achieve AYP, while their counterparts outside the urban areas sometimes outperform the paid-lunch group. Therefore, Hispanic and black students in school districts with high socioeconomic levels are far more likely to achieve AYP than their counterparts in school districts serving students from lower socioeconomic levels.

As you drill down into your own data, you gradually become aware of a paradox: as the information you acquire about the correlates of student performance in your own schools becomes more textured and valuable, the notion of measuring school performance with the types of tests mandated by NCLB becomes less plausible. That realization made it easier for me to accept the overall conclusions of my analysis. Even without including special education, thirty-two of the thirty-four Indiana Urban Schools Association school districts were ultimately identified as “Did Not Make Adequate Yearly Progress.” The special education breakouts made it unanimous.

Safe harbor or Russian roulette? To be fair, it must be noted that NCLB includes a “safe harbor” provision that schools can use to make AYP. Safe harbor is reached if the actual percentage of students failing a given test is reduced by 10 percent. Note that the percentage itself must be reduced, not the number of students. To maintain rigor where safe harbor is used to make AYP, NCLB requires that any group using AYP to reach safe harbor must also satisfy the secondary indicator. It is the only time that secondary indicators come into play for breakout groups.

At best safe harbor is a Russian roulette kind of process because, in different years, the students will be different, the “N’s” may be different, and the demographics of those tested may change. However, that is not why I downplay the safe-harbor provision here. I do so because urban school districts and schools in Indiana and most other states will miss AYP in multiple instances, leaving moot the issue of safe harbor in relation to a specific percentage.

V. Where Do We Go from Here?

Looking back, it is easy to see that my initial positive reaction to NCLB was based on general support for the act’s goals and incomplete understanding of the specifics of implementation. One early warning sign of NCLB’s deficiencies was its goal of 100 percent proficiency; nonetheless, many supporters, including this writer, generally withheld com-

ment, confident that the matter could be addressed later. *[Congress may consider changes in NCLB beginning in 2007, when the act comes up for reauthorization.—Ed.]* One hopes that such thinking does not prove wishful.

We have a saying in Indiana: “That dog won’t hunt.” Putting it bluntly, the current No Child Left Behind Act merits that description. Absent significant change, NCLB will at minimum fail to improve schooling and do nothing to aid low-performing students. In the worst-case scenario, NCLB could do serious damage to the school-improvement effort nationally and to the promising efforts already under way in states such as Indiana. The students most in need of help would be the biggest losers.

Where, then, do NCLB and AYP leave us? There is much in NCLB to like and embrace. The focus on “leaving no child behind”—systematically identifying and then addressing the needs of low-achieving students—is still a worthy if distant goal. Analyzing test scores, especially those of subgroups, is so basic to improving achievement that it is difficult to understand why its implementation required a federal mandate.

The prospects for change depend on many factors, among them political will; the willingness to face the fact that changes are needed; the staying power of those promoting change; and the true motives of those responsible for NCLB—whether they were good intentions marred by haste, disappointment over the results of past legislation, or even a desire to create a stalking-horse for vouchers. If the provisions of NCLB as they stand at this writing are not changed, the greatest consolation for the education community and all concerned may be that the results of NCLB will so lack credibility that they will be not be taken seriously. In that event, NCLB will go down as one of the greatest missed opportunities in the history of American education.

Lowell C. Rose, Ph.D., is a past executive director of Phi Delta Kappa International. Among other posts, he has also served as executive director of the Indiana School Boards Association and of the Indiana Urban Schools Association. Currently he is a consultant to school corporations on Indiana’s school improvement legislation and an adjunct professor at Indiana University.